

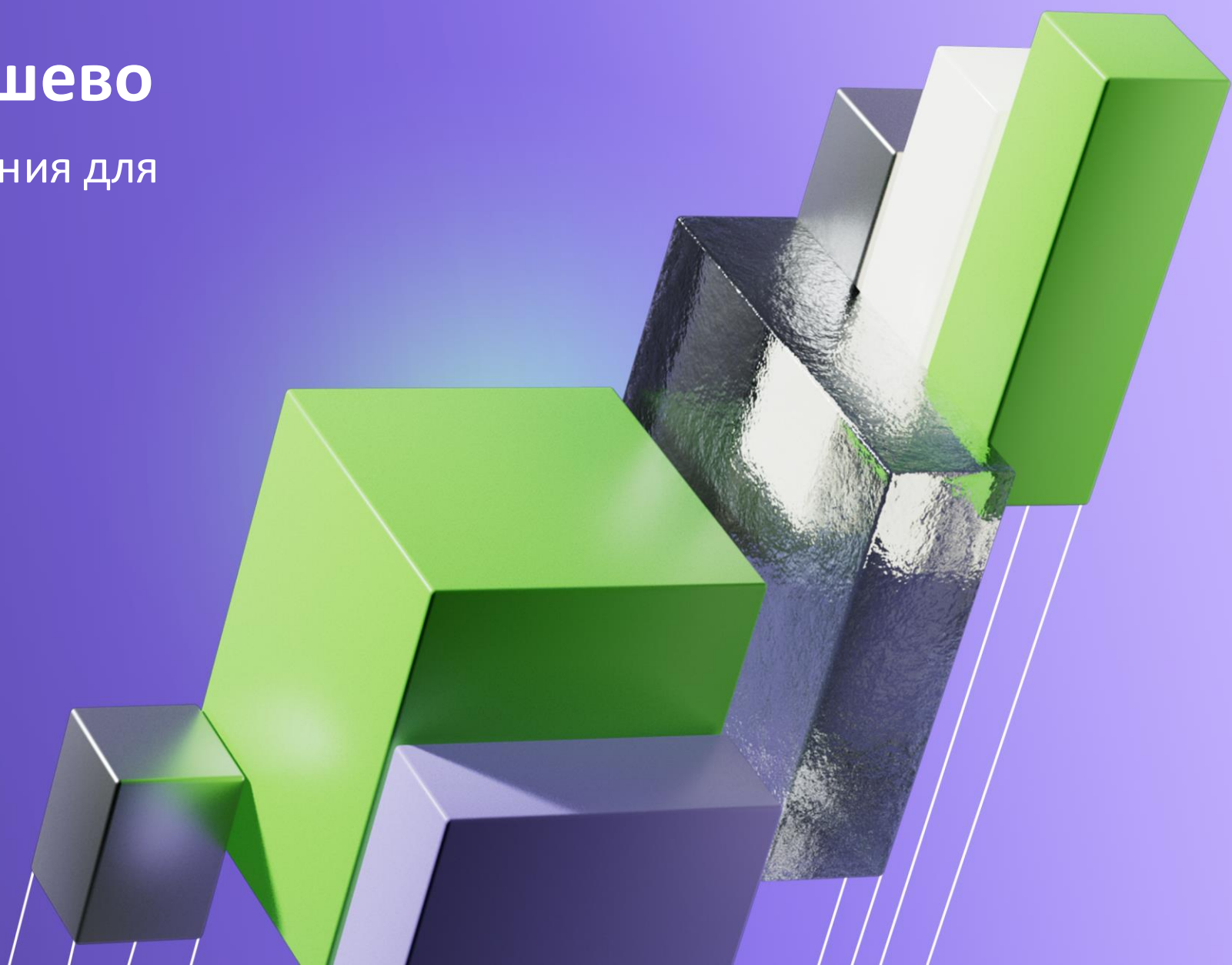
Хорошо. Быстро. Дешево

Дата-платформа нового поколения для
омни-канального Ритейлера #1



Павел Денисенко

Head of Data Architecture X5 Group



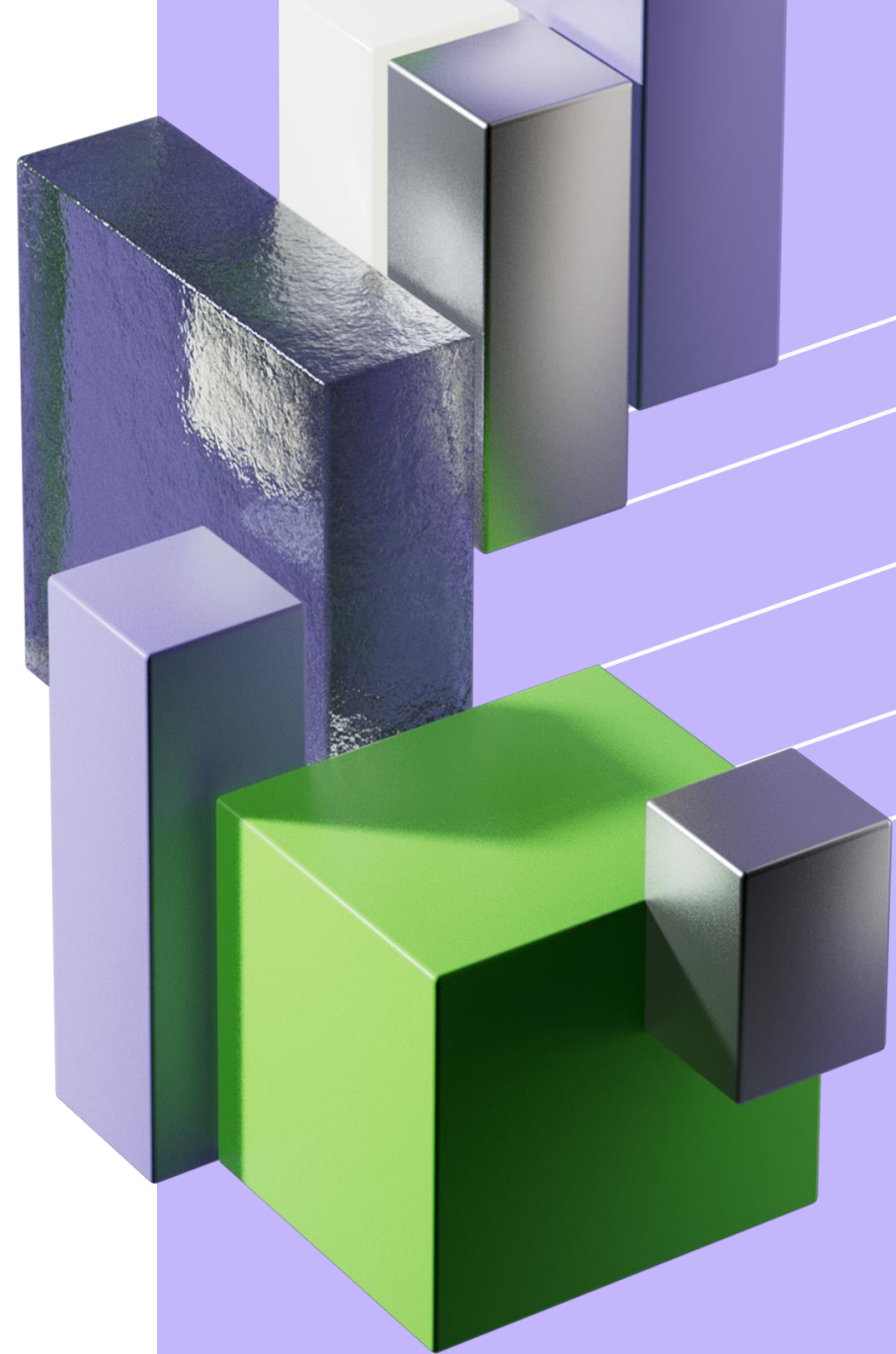


О себе в цифрах

18 

6  X5 Group

7 



X5 Group - Ритейл #1

Розница

Торговые сети

- «Пятёрочка»
- «Перекрёсток»
- «Чижик»
- Франшиза «Около»
- «Красный Яр»
- «Слата»

E-grocery

Цифровые сервисы

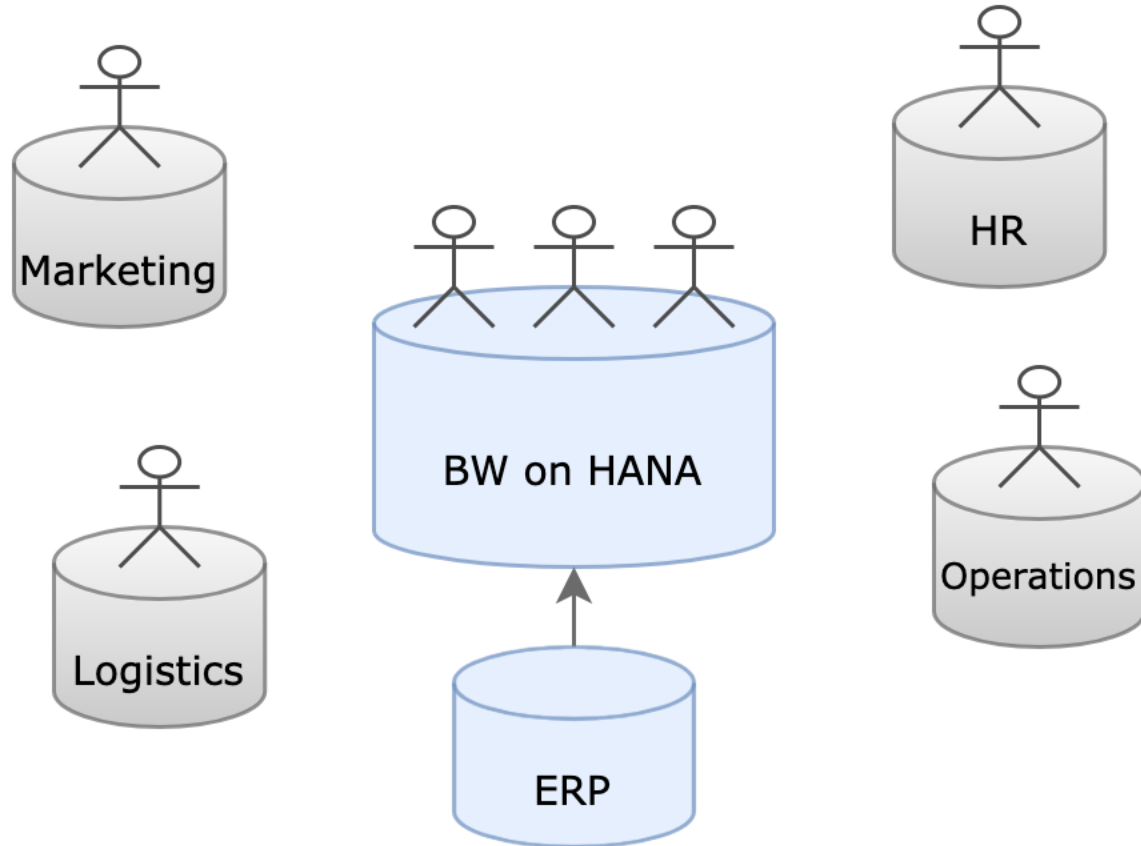
- Мобильное приложение
- Сервис «Пакет»
- Онлайн-гипермаркет «Vprok.ru»
- Онлайн-доставка
- «Много Лосося»
- 5Post
- Food.ru
- X5 Клуб

Домены

Бизнес-домены

- Операции
- Коммерция
- Клиентский опыт
- Логистика
- HR
- Финансы
- Поддержка бизнеса

2016. Ранняя федеративность



#01

У каждой бизнес-функции свое DWH

#02

Владельцы системы = пользователи

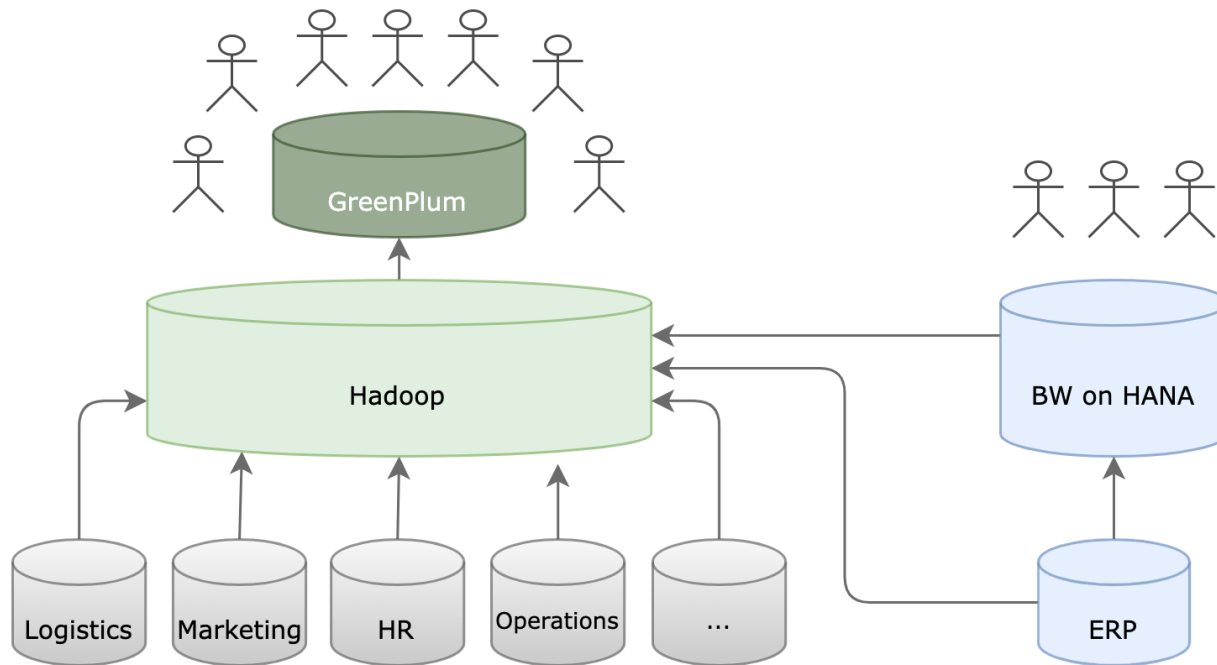
#03

Профиль потребления: BI + Excel

#04

Мощная HANA с данными из ERP, HR-источников

2019. Цифровая трансформация



#01

Data Lake = Hadoop

#02

Enterprise Data Warehouse = GreenPlum

#03

Консолидация данных из источников

#04

Единая модель данных

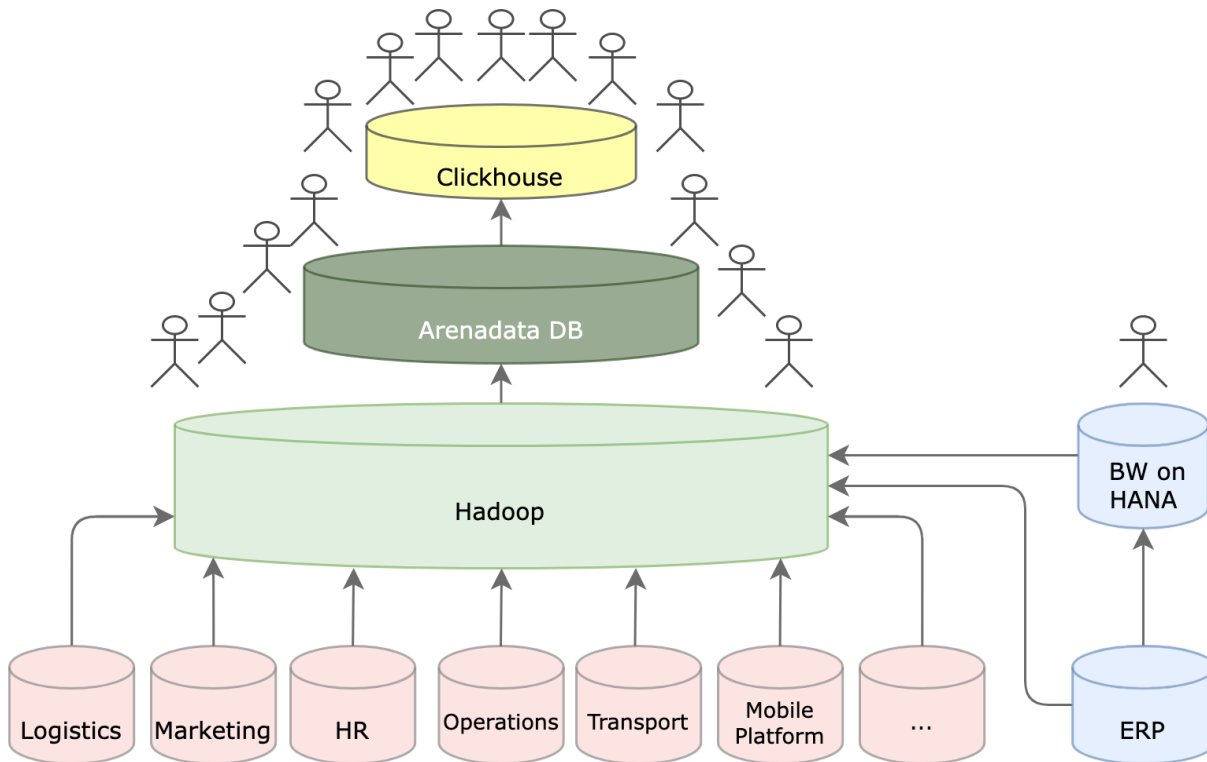
#05

Данные доступны для ML-моделей

#06

На BW – операционная отчетность

2022. Импортозамещение



#01

Clickhouse для ad-hoc highload

#02

GreenPlum от Arenadata

#03

3 экспанда кластера

#04

83 источника интегрировано

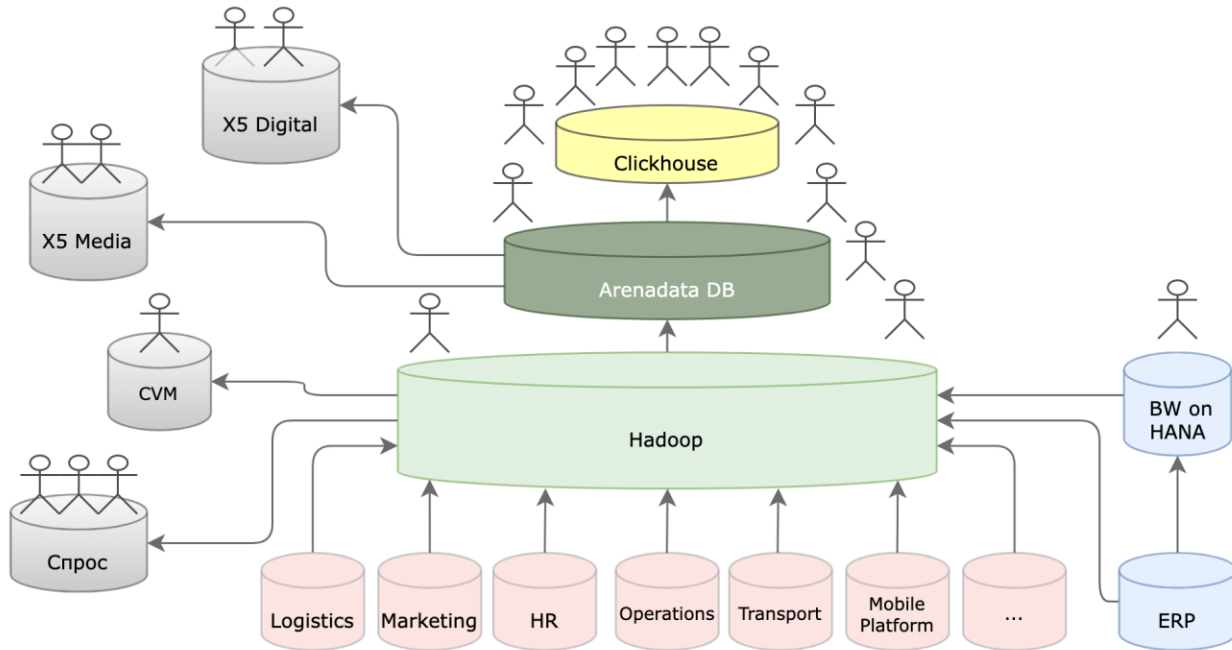
#05

Все пользователи на одной модели

#06

BW для финансового закрытия

2024. Новые бизнесы



#01

Много данных

#02

Много потребителей

#03

Конкуренция за ресурсы

#04

У каждой БЕ свои особенности

#05

У новых БЕ свое DWH

#06

Много перекладок данных



Качественные изменения бизнес модели



Экспансия в цифровой сегмент рынка



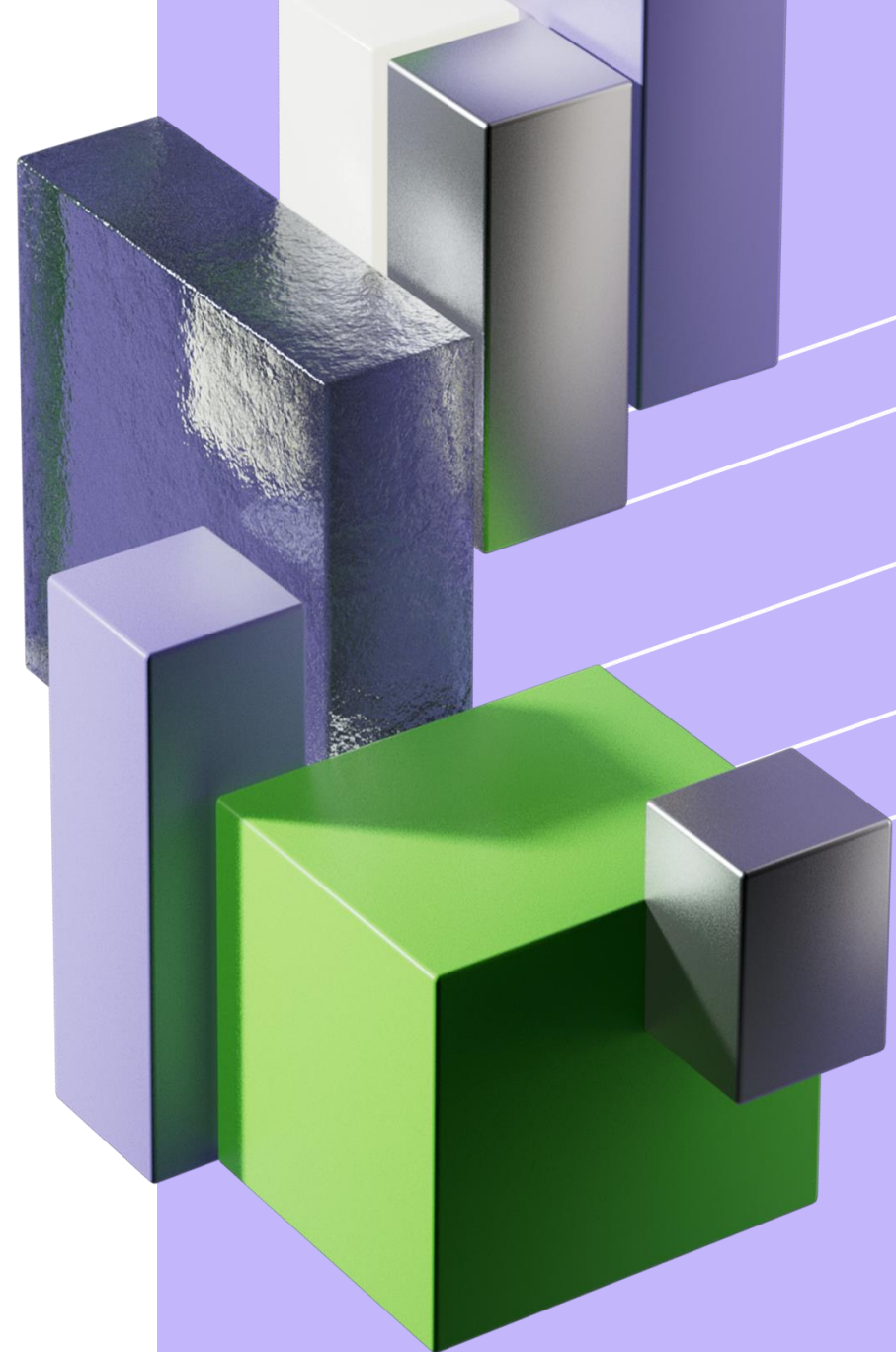
Продукты + платформы



Увеличение скорости принятия решений



Повышение эффективности затрат



Эффективная работа с данными

Скорость

- Время поиска ответа на заданный вопрос
- Срок создания нового инструмента
- Задержка в обновлении данных

Качество

- Достоверность
- Полнота
- Релевантность
- Применимость
- Прозрачность

Стоимость

- Инфраструктура
- Программное обеспечение
- Команда

Ключевые аспекты

Текущая архитектура – кратный скейлинг инфры

Гибкое управление ресурсами хранения и
вычислений, прозрачная РСМ

Независимая работа продуктовых/доменных команд

Data Mesh – принцип децентрализованного
владения данными

Разработка витрин – в продуктах, инструментах + Critical Data Elements в CDO –

Near realtime загрузки (streaming)

Хранение до 20 Пб данных

Инкремент до 100 Тб данных

Совокупный объем пользователей до 50000

X10 (на порядки) рост количества транзакций чтения
данных (до 20kk)

Время отклика < 1 мин

Новая дата-платформа: архитектура

All-in-one: все типы данных в одном storage (batch/streaming, SQL/NoSQL)

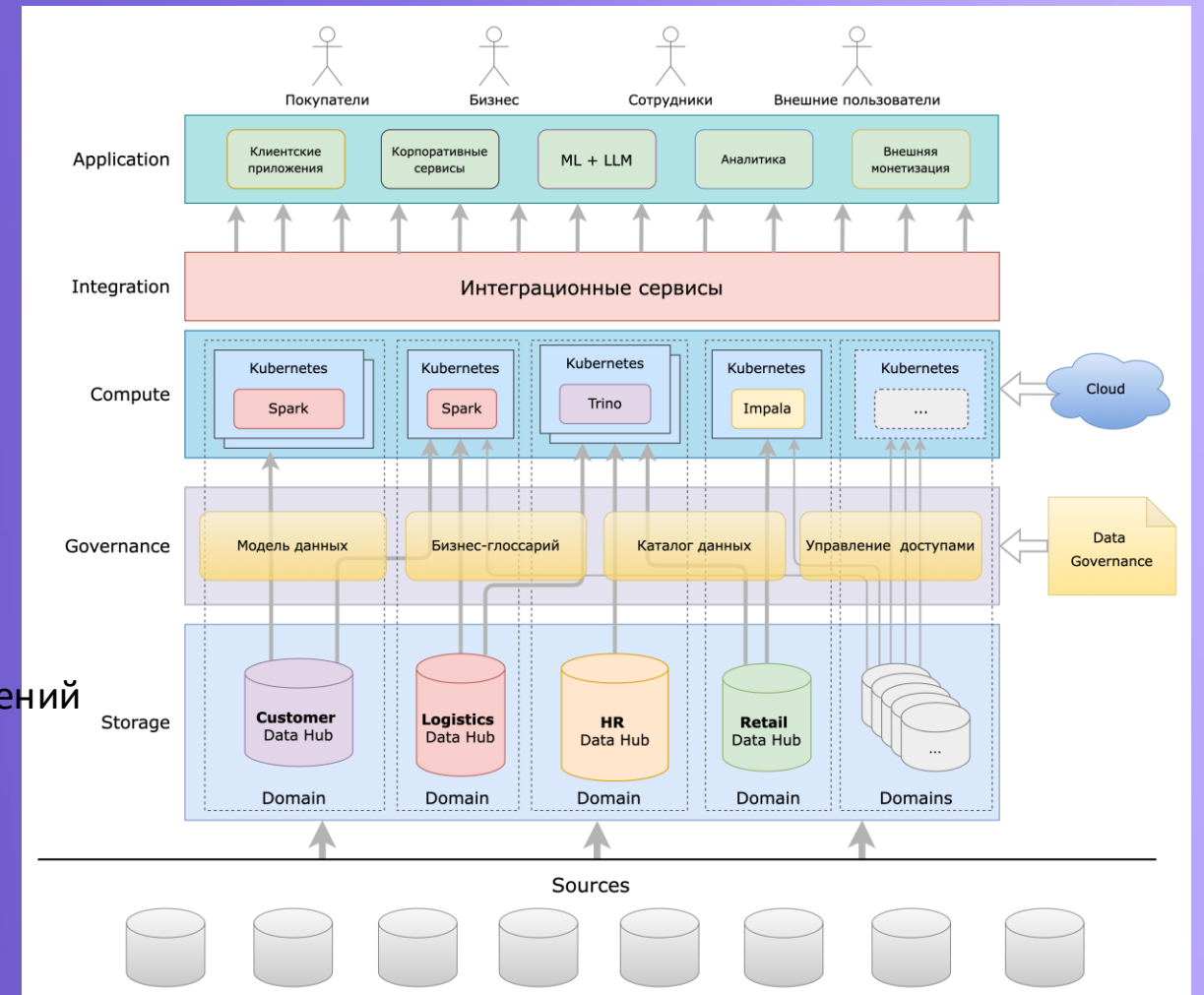
Data mesh (data hubs)

Разделение compute/storage

Composable: вариативность технологий
+ унифицированный протокол

вычислений

PaaS/DaaS – обеспечение команд гибкими инструментами
highload-работы с данными



Оптимизация времени работы магазина

На основании внешних данных о трафике и работе других объектов



до **1,5 %**
Рост доходности

Функциональность системы

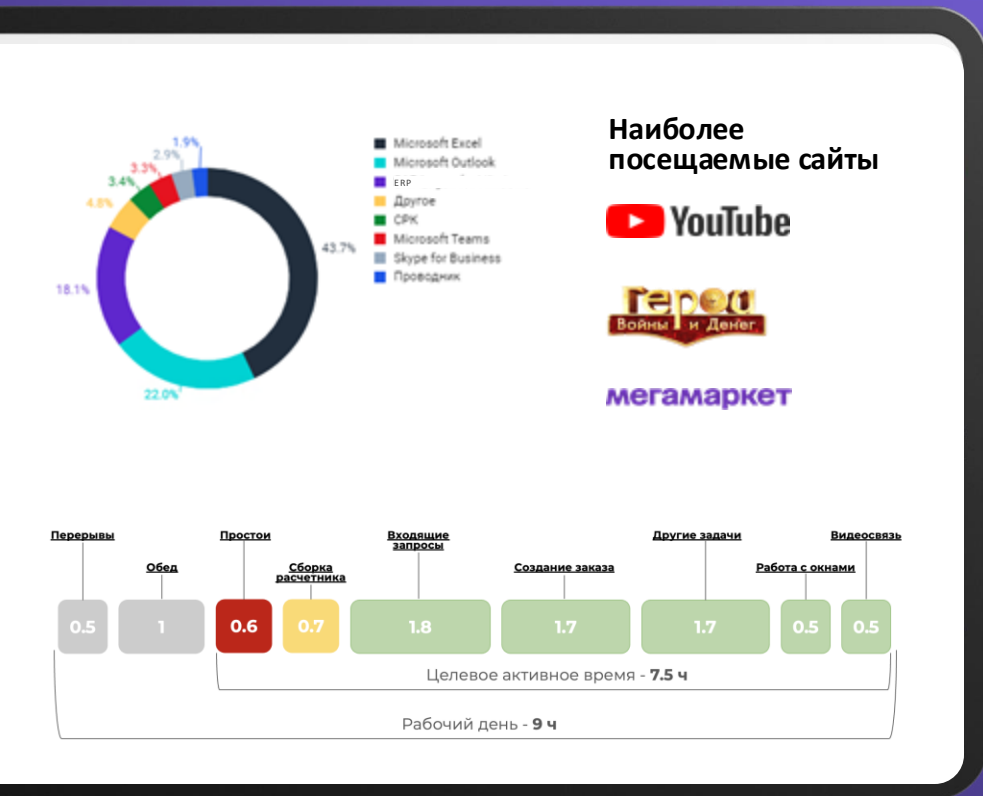
На основании данных о внешнем трафике и продажах проводится расчет конверсии

На основании стоимости работы и потенциального дохода формируется список магазинов к пересмотру времени работы

Проводится изменение времени работы магазина с учетом трафика и работы конкурентов в локации

Task Mining в Логистике

Позволяет фиксировать действия пользователя за компьютером до уровня элементарных действий



ДО **33** FTE

За счет снижения уровня простоя и непроизводительной активности

ДО **23** FTE

За счет автоматизации задач в части заполнения ручных файлов

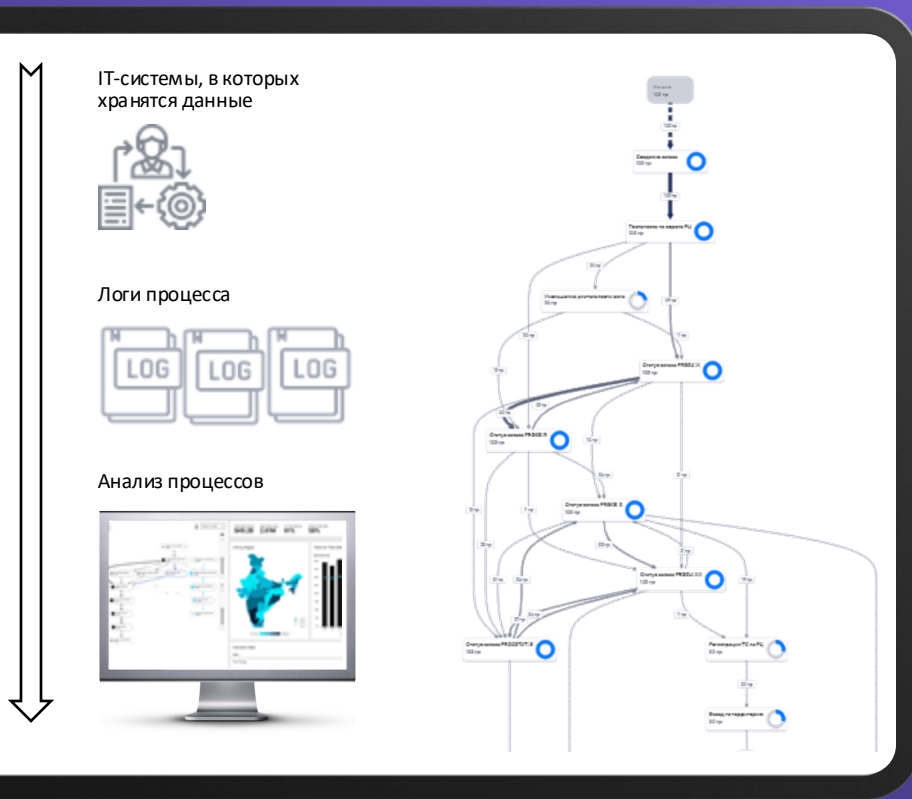
ДО **5,5** FTE

За счет RPA и внедрения промышленных решений (Qlik)



Process Mining в Логистике

Позволяет анализировать структуру и поведение процесса на основе исторических цифровых следов его прохождения в информационных системах



На **40** %
Больше палет приезжает на РЦ, чем было запланировано

2
Каждая 2-я машина приезжает раньше планового окна

60 %
Визитов дольше планового окна

Функциональность системы

Анализ цифровых следов процесса

Поиск «узкого горлышка» в процессе, точек оптимизации

Выявление отклонений от регламентированного процесса

Вопросы и ответы

